```
FILE 'HOME' ENTERED AT 09:11:56 ON 12 JUN 2002

=> file medline
COST IN U.S. DOLLARS                                    SINCE FILE      TOTAL
                                                           ENTRY    SESSION
FULL ESTIMATED COST                                         0.21       0.21

FILE 'MEDLINE' ENTERED AT 09:12:08 ON 12 JUN 2002

 FILE LAST UPDATED: 11 JUN 2002 (20020611/UP). FILE COVERS 1958 TO DATE.

 On June 9, 2002, MEDLINE was reloaded. See HELP RLOAD for details.

 MEDLINE thesauri in the /CN, /CT, and /MN fields incorporate the
 MeSH 2002 vocabulary. Enter HELP THESAURUS for details.

 THIS FILE CONTAINS CAS REGISTRY NUMBERS FOR EASY AND ACCURATE
 SUBSTANCE IDENTIFICATION.

=> s prediction (p) (coding sequences) (p) (human genes) (p) clones
         38060 PREDICTION
         17277 PREDICTIONS
         52811 PREDICTION
                 (PREDICTION OR PREDICTIONS)
         56913 CODING
            92 CODINGS
         56967 CODING
                 (CODING OR CODINGS)
        170424 SEQUENCES
          3770 CODING SEQUENCES
                 (CODING(W)SEQUENCES)
       7718955 HUMAN
        101582 HUMANS
       7738475 HUMAN
                 (HUMAN OR HUMANS)
        355134 GENES
            20 GENESES
        355153 GENES
                 (GENES OR GENESES)
          1747 HUMAN GENES
                 (HUMAN(W)GENES)
         53317 CLONES
L1          27 PREDICTION (P) (CODING SEQUENCES) (P) (HUMAN GENES) (P) CLONES

=> s l1 (p) cDNA (p) brain
         81145 CDNA
         15300 CDNAS
         88437 CDNA
                 (CDNA OR CDNAS)
        544478 BRAIN
         20866 BRAINS
        548623 BRAIN
                 (BRAIN OR BRAINS)
L2          17 L1 (P) CDNA (P) BRAIN

=> s l2 (p) (large proteines)
        386015 LARGE
            39 LARGES
        386039 LARGE
                 (LARGE OR LARGES)
          1563 PROTEINES
             0 LARGE PROTEINES
                 (LARGE(W)PROTEINES)
L3           0 L2 (P) (LARGE PROTEINES)

=> s l2 (p) (in vitro)
        694398 VITRO
            49 VITROS
        694444 IN VITRO
                 (VITRO OR VITROS)
L4          15 L2 (P) (IN VITRO)
```

```
=> s l4 (p) proteins
        901571 PROTEINS
L5          15 L4 (P) PROTEINS

=> s l4 (p) (large proteins)
        386015 LARGE
            39 LARGES
        386039 LARGE
                  (LARGE OR LARGES)
        901571 PROTEINS
           377 LARGE PROTEINS
                  (LARGE(W)PROTEINS)
L6          15 L4 (P) (LARGE PROTEINS)

=> s l6 and py=1998
        442735 PY=1998
L7           5 L6 AND PY=1998

=> d l7 1-5 ibib abs

L7   ANSWER 1 OF 5     MEDLINE
ACCESSION NUMBER:   1999156230     MEDLINE
DOCUMENT NUMBER:    99156230   PubMed ID: 10048485
TITLE:                ***Prediction***   of the   ***coding***
                      ***sequences***   of unidentified   ***human***
                      ***genes***  . XII. The complete sequences of 100 new
                      ***cDNA***     ***clones***   from   ***brain***   which
                    code for   ***large***     ***proteins***   in
                      ***vitro***  .
AUTHOR:             Nagase T; Ishikawa K; Suyama M; Kikuno R; Hirosawa M;
                    Miyajima N; Tanaka A; Kotani H; Nomura N; Ohara O
CORPORATE SOURCE:   Kazusa DNA Research Institute, Kisarazu, Chiba, Japan.
SOURCE:             DNA RESEARCH,   ***(1998 Dec 31)***   5 (6) 355-64.
                    Journal code: 9423827. ISSN: 1340-2838.
PUB. COUNTRY:       Japan
                    Journal; Article; (JOURNAL ARTICLE)
LANGUAGE:           English
FILE SEGMENT:       Priority Journals
OTHER SOURCE:       GENBANK-AB020626; GENBANK-AB020627; GENBANK-AB020628;
                    GENBANK-AB020629; GENBANK-AB020630; GENBANK-AB020631;
                    GENBANK-AB020632; GENBANK-AB020633; GENBANK-AB020634;
                    GENBANK-AB020635; GENBANK-AB020636; GENBANK-AB020637;
                    GENBANK-AB020638; GENBANK-AB020639; GENBANK-AB020640;
                    GENBANK-AB020641; GENBANK-AB020642; GENBANK-AB020643;
                    GENBANK-AB020644; GENBANK-AB020645; GENBANK-AB020646;
                    GENBANK-AB020647; GENBANK-AB020648; GENBANK-AB020649;
                    GENBANK-AB020650; GENBANK-AB020651; GENBANK-AB020652;
                    GENBANK-AB020653; GENBANK-AB020654; GENBANK-AB020655
ENTRY MONTH:        199905
ENTRY DATE:         Entered STN: 19990517
                    Last Updated on STN: 20000303
                    Entered Medline: 19990506
AB  In this paper, we report the sequences of 100 cDNA clones newly determined
    from a set of size-fractionated human brain cDNA libraries and predict the
    coding sequences of the corresponding genes, named KIAA0819 to KIAA0918.
    These cDNA clones were selected on the basis of their coding potentials of
    large proteins (50 kDa and more) by using in vitro
    transcription/translation assays. The sequence data showed that the
    average sizes of the inserts and corresponding open reading frames are 4.4
    kb and 2.5 kb (831 amino acid residues), respectively. Homology and
    motif/domain searches against the public databases indicated that the
    predicted coding sequences of 83 genes were similar to those of known
    genes, 59% of which (49 genes) were categorized as coding for proteins
    functionally related to cell signaling/communication, cell
    structure/motility and nucleic acid management. The chromosomal locations
    and the expression profiles of all the genes were also examined. For 54
    clones including brain-specific ones, the mRNA levels were further
    examined among 8 brain regions (amygdala, corpus callosum, cerebellum,
    caudate nucleus, hippocampus, substantia nigra, subthalamic nucleus, and
    thalamus), spinal cord, and fetal brain.
```

L7 ANSWER 2 OF 5 MEDLINE
ACCESSION NUMBER: 1999087487 MEDLINE
DOCUMENT NUMBER: 99087487 PubMed ID: 9872452
TITLE: ***Prediction*** of the ***coding*** ***sequences*** of unidentified ***human*** ***genes*** . XI. The complete sequences of 100 new ***cDNA*** ***clones*** from ***brain*** which code for ***large*** ***proteins*** in ***vitro*** .
AUTHOR: Nagase T; Ishikawa K; Suyama M; Kikuno R; Miyajima N; Tanaka A; Kotani H; Nomura N; Ohara O
CORPORATE SOURCE: Kazusa DNA Research Institute, Kisarazu, Chiba, Japan.
SOURCE: DNA RESEARCH, ***(1998 Oct 30)*** 5 (5) 277-86.
Journal code: 9423827. ISSN: 1340-2838.
PUB. COUNTRY: Japan
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AB018254; GENBANK-AB018255; GENBANK-AB018256; GENBANK-AB018257; GENBANK-AB018258; GENBANK-AB018259; GENBANK-AB018260; GENBANK-AB018261; GENBANK-AB018262; GENBANK-AB018263; GENBANK-AB018264; GENBANK-AB018265; GENBANK-AB018266; GENBANK-AB018267; GENBANK-AB018268; GENBANK-AB018269; GENBANK-AB018270; GENBANK-AB018271; GENBANK-AB018272; GENBANK-AB018273; GENBANK-AB018274; GENBANK-AB018275; GENBANK-AB018276; GENBANK-AB018277; GENBANK-AB018278; GENBANK-AB018279; GENBANK-AB018280; GENBANK-AB018281; GENBANK-AB018282; GENBANK-AB018283; +
ENTRY MONTH: 199904
ENTRY DATE: Entered STN: 19990504
Last Updated on STN: 19990504
Entered Medline: 19990419
AB In our series of projects for accumulating sequence information on the coding sequences of unidentified human genes, we have newly determined the sequences of 100 cDNA clones from a set of size-fractionated human brain cDNA libraries, and predicted the coding sequences of the corresponding genes, named KIAA0711 to KIAA0810. These cDNA clones were selected according to their coding potentials of large proteins (50 kDa and more) in vitro. The average sizes of the inserts and corresponding open reading frames were 4.3 kb and 2.6 kb (869 amino acid residues), respectively. Sequence analyses against the public databases indicated that the predicted coding sequences of 78 genes were similar to those of known genes, 64% of which (50 genes) were categorized as proteins functionally related to cell signaling/communication, cell structure/motility and nucleic acid management. As additional information concerning genes characterized in this study, the chromosomal locations of the clones were determined by using human-rodent hybrid panels and the expression profiles among 10 human tissues were examined by reverse transcription-coupled polymerase chain reaction which was substantially improved by enzyme-linked immunosorbent assay.

L7 ANSWER 3 OF 5 MEDLINE
ACCESSION NUMBER: 1998403880 MEDLINE
DOCUMENT NUMBER: 98403880 PubMed ID: 9734811
TITLE: ***Prediction*** of the ***coding*** ***sequences*** of unidentified ***human*** ***genes*** . X. The complete sequences of 100 new ***cDNA*** ***clones*** from ***brain*** which can code for ***large*** ***proteins*** in ***vitro*** .
AUTHOR: Ishikawa K; Nagase T; Suyama M; Miyajima N; Tanaka A; Kotani H; Nomura N; Ohara O
CORPORATE SOURCE: Kazusa DNA Research Institute, Kisarazu, Chiba, Japan.
SOURCE: DNA RESEARCH, ***(1998 Jun 30)*** 5 (3) 169-76.
Journal code: 9423827. ISSN: 1340-2838.
PUB. COUNTRY: Japan
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AB014511; GENBANK-AB014512; GENBANK-AB014513; GENBANK-AB014514; GENBANK-AB014515; GENBANK-AB014516; GENBANK-AB014517; GENBANK-AB014518; GENBANK-AB014519;

GENBANK-AB014[redacted]; GENBANK-AB014521; GENBANK-AB014522;
GENBANK-AB014[redacted]; GENBANK-AB014524; GENBANK-AB014525;
GENBANK-AB014526; GENBANK-AB014527; GENBANK-AB014528;
GENBANK-AB014529; GENBANK-AB014530; GENBANK-AB014531;
GENBANK-AB014532; GENBANK-AB014533; GENBANK-AB014534;
GENBANK-AB014535; GENBANK-AB014536; GENBANK-AB014537;
GENBANK-AB014538; GENBANK-AB014539; GENBANK-AB014540; +
ENTRY MONTH: 199812
ENTRY DATE: Entered STN: 19990115
Last Updated on STN: 19990115
Entered Medline: 19981209

AB As an extension of our cDNA analysis for deducing the coding sequences of unidentified human genes, we have newly determined the sequences of 100 cDNA clones from a set of size-fractionated human brain cDNA libraries, and predicted the coding sequences of the corresponding genes, named KIAA0611 to KIAA0710. In vitro transcription-coupled translation assay was applied as the first screening to select cDNA clones which produce proteins with apparent molecular mass of 50 kDa and over. One hundred unidentified cDNA clones thus selected were then subjected to sequencing of entire inserts. The average size of the inserts and corresponding open reading frames was 4.9 kb and 2.8 kb (922 amino acid residues), respectively. Computer search of the sequences against the public databases indicated that predicted coding sequences of 87 genes were similar to those of known genes, 62% of which (54 genes) were categorized as proteins related to cell signaling/communication, cell structure/motility and nucleic acid management. The expression profiles in 10 human tissues of all the clones characterized in this study were examined by reverse transcription-coupled polymerase chain reaction and the chromosomal locations of the clones were determined by using human-rodent hybrid panels.

L7 ANSWER 4 OF 5 MEDLINE
ACCESSION NUMBER: 1998340016 MEDLINE
DOCUMENT NUMBER: 98340016 PubMed ID: 9675416
TITLE: Characterization of a new beta-spectrin gene which is predominantly expressed in brain.
AUTHOR: Ohara O; Ohara R; Yamakawa H; Nakajima D; Nakayama M
CORPORATE SOURCE: Laboratory of DNA Technology, Kazusa DNA Research Institute, Chiba, Japan.. ohara@kazusa.or.jp
SOURCE: BRAIN RESEARCH. MOLECULAR BRAIN RESEARCH, ***(1998 Jun*** *** 15)*** 57 (2) 181-92.
Journal code: 8908640. ISSN: 0169-328X.
PUB. COUNTRY: Netherlands
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AB008551; GENBANK-AB008567
ENTRY MONTH: 199810
ENTRY DATE: Entered STN: 19981029
Last Updated on STN: 20000303
Entered Medline: 19981020

AB We recently identified a gene which shows high similarity to the beta-spectrin gene but with a different chromosomal location from either of the two known beta-spectrin genes [T. Nagase, K.-I. Ishikawa, D. Nakajima, M. Ohira, N. Seki, N. Miyajima, A. Tanaka, H. Kotani, N. Nomura, O. Ohara, ***Prediction*** of the ***coding*** ***sequences*** of unidentified ***human*** ***genes*** : VII. The complete sequences of 100 new ***cDNA*** ***clones*** from ***brain*** which can code for ***large*** ***proteins*** in ***vitro*** , DNA Res. 4 (1997) 141-150]. In order to further characterize this new spectrin gene and its product, we isolated the rat counterpart of this gene and analyzed it in terms of its protein coding sequence, the tissue distribution of its mRNA and the product, and the regional distribution of the mRNA and the product in the ***brain*** . The results indicated that this gene was most abundantly transcribed in the ***brain*** and neurons were the predominant cell-type to express this gene. In particular, Purkinje cells were the richest in this gene product, and this new form of beta-spectrin was found more prominently in the dendrites than in the cell bodies. Since the expression pattern and the subcellular localization of this gene product were quiet distinct from those of the two beta-spectrin isoforms already characterized, this beta-spectrin gene would play an important role in neuronal membrane skeleton although it has

been overlooked to date.

L7 ANSWER 5 OF 5 MEDLINE
ACCESSION NUMBER: 1998290545 MEDLINE
DOCUMENT NUMBER: 98290545 PubMed ID: 9628581
TITLE: ***Prediction*** of the ***coding*** ***sequences*** of unidentified ***human*** ***genes*** . IX. The complete sequences of 100 new ***cDNA*** ***clones*** from ***brain*** which can code for ***large*** ***proteins*** in ***vitro*** .
AUTHOR: Nagase T; Ishikawa K; Miyajima N; Tanaka A; Kotani H; Nomura N; Ohara O
CORPORATE SOURCE: Kazusa DNA Research Institute, Kisarazu, Chiba, Japan.
SOURCE: DNA RESEARCH, ***(1998 Feb 28)*** 5 (1) 31-9.
Journal code: 9423827. ISSN: 1340-2838.
PUB. COUNTRY: Japan
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AB011083; GENBANK-AB011084; GENBANK-AB011085; GENBANK-AB011086; GENBANK-AB011087; GENBANK-AB011088; GENBANK-AB011089; GENBANK-AB011090; GENBANK-AB011091; GENBANK-AB011092; GENBANK-AB011093; GENBANK-AB011094; GENBANK-AB011095; GENBANK-AB011096; GENBANK-AB011097; GENBANK-AB011098; GENBANK-AB011099; GENBANK-AB011100; GENBANK-AB011101; GENBANK-AB011102; GENBANK-AB011103; GENBANK-AB011104; GENBANK-AB011105; GENBANK-AB011106; GENBANK-AB011107; GENBANK-AB011108; GENBANK-AB011109; GENBANK-AB011110; GENBANK-AB011111; GENBANK-AB011112
ENTRY MONTH: 199808
ENTRY DATE: Entered STN: 19980903
Last Updated on STN: 20000303
Entered Medline: 19980825

AB As an extension of a series of projects for sequencing human cDNA clones derived from relatively long transcripts, we herein report the entire sequences of 100 newly determined cDNA clones with the potential of coding for large proteins in vitro. The cDNA clones were isolated from size-fractionated human brain cDNA libraries with insert sizes between 4.5 and 8.3 kb. The sequencing of these clones revealed that the average size of the cDNA inserts and of their open reading frames was 5.3 kb and 2.8 kb (930 amino acid residues), respectively. Homology search against public databases indicated that the predicted coding sequences of 86 clones exhibited significant similarities to known genes; 51 of them (59%) were related to those for cell signaling/communication, nucleic acid management, and cell structure/motility. All the clones characterized in this study are accompanied by their expression profiles in 14 human tissues examined by reverse transcription-coupled polymerase chain reaction and the chromosomal mapping data.

=> log y
COST IN U.S. DOLLARS

	SINCE FILE ENTRY	TOTAL SESSION
FULL ESTIMATED COST	4.04	4.25

STN INTERNATIONAL LOGOFF AT 09:17:08 ON 12 JUN 2002